X-KEYSCORE
Application IDs

xkeyscore@nsa

Basic Syntax

Syntax:

- Very C-like.

function('name', level, <optional info>) = 'search terms and pattern';

Two valid search functions appid and fingerprint:

- *appid('chat/icq', 8.5, wireshark='icq', chatproc='ICQ') =*
  /[^o]icq/c and $icq;

- *fingerprint('fingerprint/phone/nokia/generic', 7.0) =*
  'user-agent: nokia' or
  'profile: http://nds.nokia.com/uaprof/n';

Naming Conventions

XKS Appid's are named using a pseudo directory convention.

/application_type/sub_type/name

Levels

Levels are 1.0 – 9.9 with lower numbers having a higher priority. This allows multiple signatures to match a piece of traffic and only the most specific appid will be applied. An example might be:

9.9 Yahoo

9.8 Yahoo/chat

9.7 Yahoo/chat/incoming

Since the Yahoo/chat/incoming has the lowest level, the traffic will be labeled as yahoo/chat/incoming

Basic Search Patterns

XKEYSCORE supports Boolean operations and regular expressions

Raw text must be encapsulated between single quotes

- *'search term'*

Terms can be combined with Boolean logic

- *'search term' and 'another term'*
- *'search term' or 'another term'*

Binary and Regex Patterns

Binary patterns can be represented by putting a \x in front of each binary value

- '*\xff\xff\x00\x02*'

Note: Unlike C, no double back slashing required

/regex/

CHAINWORDs

You can assign a pattern to a variable (CHAINWORD) and reuse the variable in many patterns.

- *$sip = 'via: sip' and 'cseq:' and 'SIP/2'c;*

Now we can use this variable in future definitions:

- *appid('voip/sip', 7.2) = $sip;*

- *appid('voip/sip/invite', 6.9) = $sip and 'INVITE';*

Built in functions

ip(expr)	Matches against an IP Address looks in to address and from address in the session headere • ip('10.10.10.1');
toport(expr)	Matches against the Destination/To port. Note this must be a numeric representation of a port. • toport(1920);
fromport(expr)	Matches against the Source/From port. Note this must be a numeric representation of a port. • fromport(80);
port(expr)	Matches against the either port. Note this must be a numeric representation of a port. •port(6667);
next_protocol(expr)	Matches against the integer version of the next protocol. • next_protocol(250);
protocol ('text')	Will only work for IP next protocol names as defined in the IANA next protocol numbers document • protocol('tcp');

Built in functions

email_address(sel)	permutes just like strong_selector (just like DECODEORDAIN
mac_address(addr)	Tasks a mac address
smac(addr)	
dmac(addr)	
ip(addr)	tasks this IP address (either to or from)
from_ip(addr)	tasks this IP address only when it is the originator
to_ip(addr)	tasks this IP address only when it is the destination

More built in functions

first(expr)	Matches against a pattern at the beginning of the session
lpos(expr)	Matches against a pattern at the beginning of each line (\n)
pos(expr)	expression occurs at offset X in the session • pos('Hello') == 5, • pos(/Good.*Grief/) <= 10
between(expr)	• between('Hello', 'World', 10, 100) Separation between 'Hello' and 'World' is greater than or equal to 10 bytes and less than or equal to 100 bytes This is the same as using the following regular expression: • /Hello.{10,100}World/
'term'c	Does a case sensitive match of the term
'term'u	Treats the term as UTF-16

Predefined Chainwords

There are a number of chainwords predefined for convenience:

- $tcp
- $udp
- $icmp
- $sctp
- $rpc
- $arp
- $ssl
- $http_cmd
- $http
- $http_get
- $http_put
- $http_post
- $http_delete
- $http_trace
- $http_head
- $http_options
- $http_partial
- $vbulletin
- $mime_type
- $user_agent

Example

```
appid('voip/sip/IMS', 6.0, wireshark='sip') =
                ('via: sip' or 'v: sip') and 'cseq:' and (
                  'p-access-network-info:' or
                  'p-called-party-id:' or
                  'p-charging-vector:' or
                  'p-charging-vector-addresses:' or
                  'p-media-authorization:' or
                  'security-verify:' or
                  'security-server:' or
                  'security-client:' or
                  'service-route:' or
                  'record-route:' and 'pcscf' or
                  'record-route:' and 'scscf' or
                  'contact:' and 'pcscf' or
                  'contact:' and 'scscf' or
                  'proxy-authorization:' and 'pcscf' or
                  'proxy-authorization:' and 'scscf' or
                  'path:' and 'pcscf' or
                  'path:' and 'scscf'
                );
```

Example

```
appid('voip/skinny/keep-alive', 3.0, wireshark='skinny') =
                toport(2000) and
    first('\x04\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00')
    ;

appid('voip/skinny/keep-alive-ack', 3.0, wireshark='skinny') =
                fromport(2000) and
    first('\x04\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00')
    ;

appid('voip/skinny(port2000)', 9.9, wireshark='skinny') =
                port(2000);
```

Example

```
appid('chat/yahoo', 6.0, chatproc='Yahoo') =
            (('YCHT'c and $yahoo_chat) or first('YCHT'c)) and not
    port(5050);

appid('chat/icq', 8.5, wireshark='icq', chatproc='ICQ') =
            /[^o]icq/c and $icq;


appid('chat/icq', 9.0, wireshark='icq', chatproc='ICQ') =
            first('icq') and not port(25);

fingerprint('encryption/moujahedeen', 7.0) =
            'begin+gimf+asrar+el+moujahedeen' or
            'begin gimf asrar el moujahedeen';
```

Example

```
appid('mail/smtp/to_server', 8.5, direction=$from_server,
    wireshark='smtp') =
                toport(25) and
                ( first('helo') or
                  first('ehlo') or
                first('data') or
                (lpos('To: 'c) and lpos('From:'c)) or
                lpos('QUIT'c) or
                lpos('mail from:') or
                lpos('rcpt to:') );
```

Example

```
$gmail =        '<script>D=(top.js&&top.js.init)?function(d){top.js.P(window'c or
              first('POST /gmail'c) or
              first('GET /gmail'c) or
              'GMAIL_AT='c or
              /SID=[A-Za-z0-9\-\_]{87}=;Domain=\.google\.com/c or
              'GMAIL_STAT='c or
              '[[\"ct\"'c or
              'S=gmail='c or
              'ain=\'mail.google.com'c or
              '<title>Gmail'c or
              'GMAIL_RTT='c or
              'GMAIL_LOGIN='c or
              '\nServer: GFE/'c;
appid('mail/webmail/gmail', 8.0, webproc='Gmail') =
              $gmail;
```

Append Option

```
# append the mime_type and HTML title to any of these appids..
PARAMS = append=$mime_type, append2=$http_info,
    append3=$doc_title;

$web = "web";

appid('http/proxy_to_server', 9.1, $web, direction=$proxy_to_server) =
            $webproxy_to_server;

appid('http/proxy_to_client', 9.1, $web, direction=$proxy_to_client) =
            $webproxy_to_client;
```

Type Option

Third parameter is the type; if missing, it takes up to the first slash as the type

```
appid('http/response', 9.2, $web) =
        $http and
        not ('x-cache' or 'x-forward' or 'get /' or
                'post /' or 'get http' or 'post http');

appid('http/response/partial', 9.1, $web) =
        $http and $http_partial;
```

Appid utility

```
appid options:
  --help                  this help message
  --list-all              list all the application/fingerprint names and
                          levels
  --list-appids           list all the application names (no fingerprints)
  --list-fingerprints     list all the application names (no appids)
  --list-types            list all the application types
  --list-levels           list all the application levels
  --unit-test             perform unit tests with data in the heirachy
                          'datadir', with files matching 'filespec'
  --quiet                 don't print any load messages
  --appid_fname arg       location of appid.cfg
  --input-file arg        input file to test
  --datadir arg           The test data directory. Defaults to
                          $(XSCORE_TEST_DATA_DIR)/appids
  --filespec arg (=.*\.u124) A regular expression to match against files to
                          check
  --noexit arg (=0)       do not stop on the first error
```

Appid Validation

```
appid sample.u124
Loading appids
->Loading : /home/oper/xkeyscore/config/dictionaries/appid/appid_definitions.cfg
->Loading : /home/oper/xkeyscore/config/dictionaries/appid/anonymizer.appid
->Loading : /home/oper/xkeyscore/config/dictionaries/appid/bulletin_board.appid
->Loading : /home/oper/xkeyscore/config/dictionaries/appid/tao_vpn.appid
->Loading : /home/oper/xkeyscore/config/dictionaries/appid/tdmoip.appid
->Loading : /home/oper/xkeyscore/config/dictionaries/appid/terminal.appid
->Loading : /home/oper/xkeyscore/config/dictionaries/appid/voip.appid
->Loading : /home/oper/xkeyscore/config/dictionaries/appid/appid_definitions.cfg
Finished loading appids
Filename: sample.u124

Appid: encryption/https

Total Size: 19.36Kbits
Total Time: 0.01secs
Rate: 1.936Mbits/s
Overall performance:
Total Time: 0.01secs
Total Bits: 0.01936Mbits
Overall Rate: 1.936Mbits/s
```

```
<entry>
 <type> boolean </type>
 <name> chat/irc </name>
 <category> application_i
 <information>
  <application type> chat
  <level> 8.1 </level>
  <action> application_id
 </information>
 <select>
  KW='<POS=0>nick'+
  KW='<POS=0>ison'+
  KW='<POS=0>whois'+
  KW='<POS=0>isirc'+
  KW='<POS=0>irc'+
  KW='<POS=0>join'+
  KW='<POS=0>auth'+
  KW='<POS=0>crypt'+
  KW='<POS=0>ping'+
  KW='<POS=0>pong'+
  KW='<POS=0>privmsg'+
  KW='<POS=0>notice auth'+
  KW='irc'&KW='privmsg'&KW='notice'+
  KW='error :you'
 </select>
 <deselect> KW='ap 0013'&KW='user' </deselect>
</entry>
```

```
appid('chat/irc', 8.5, wireshark='irc', chatproc='IRC') =
                'privmsg ';

appid('chat/irc', 8.1, wireshark='irc', chatproc='IRC') =
    not (port(113) and 'user') and
    ( first('nick ') or first('ison ') or
      first('whois ') or first('isirc ') or
      first('irc ')  or first('join ') or
      first('auth ') or first('crypt ') or
      first('ping ') or first('pong ') or
      first('privmsg ') or
      first('notice auth') or
      ('irc' and 'privmsg ' and 'notice ')
      or 'error :you'
    );
```